Chapter 3 Markov chains for the long term $\boxed{\text{R demo}}$

Section 3.1 Limiting distribution

Def'n: Let $X_0, X_1, X_2, \cdots$ be a Markov chain with transition matrix P. A <u>limiting distribution</u> for the Markov chain is a probability distribution $\bar{\lambda}$ with the property that for all states i and j,

$$\lim_{n\to\infty} P^n_{ij} = \lambda_j \quad (ij^{th} \text{ entry of } P^n \text{ equals } \lambda_j)$$

Equivalently, $\lim_{n\to\infty} \alpha P^n = \bar{\lambda}$ for all initial prob vectors α

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \bar{\lambda} \\ \vdots \\ \bar{\lambda} \end{bmatrix} \quad (\text{all rows are copies of } \lambda)$$

Interpretation: λ_j gives proportion of time in state j over long run

$\lambda_j \approx$ prob $X_n = j$ for large n

Section 3.2 Stationary distribution

Def'n: Let $X_0, X_1, X_2, \cdots$ be a Markov chain with transition matrix P.
A <u>stationary distribution</u> is a prob distribution $\bar{\pi}$ that satisfies $\bar{\pi} = \bar{\pi} P$.

↳ also called <u>invariant</u> vector

That is, $\pi_j = \sum_i \pi_i P_{ij}$ for all states j.

Interpretation: If the current prob distr of states is $\bar{\pi}$, it will remain $\bar{\pi}$ for all future steps – each time we multiply by P, we get $\bar{\pi}$ back.

Suppose the initial prob vector is $\bar{\pi}$ (for X_0)
The prob vector for X_1 is then $\bar{\pi} P = \bar{\pi}$

X_2 $\quad \bar{\pi} P^2 = \bar{\pi} P = \bar{\pi}$

$\vdots$

X_n $\quad \bar{\pi} P^n = \bar{\pi}$

Question: How are limiting & stationary distributions related?

Answer: <u>Lemma 3.1</u> Limiting distributions are stationary distributions.

Proof: Suppose $\bar{\pi}$ is a limiting distribution:

$$\lim_{n\to\infty} (P^n)_{ij} = \pi_j \text{ for all states } i \text{ and } j$$

Need to show $\bar{\pi} = \bar{\pi} P$.

$$\pi_j = \lim_{n\to\infty} P^n_{ij} = \lim_{n\to\infty} P^{n+1}_{ij} = \lim_{n\to\infty} \sum_k \underbrace{P^n_{ik}}\, P_{kj}$$

$$= \sum_k \pi_k P_{kj}$$

$\Rightarrow \bar{\pi} = \bar{\pi} P$. $\square$

Converse is not true:

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$\bar{\pi} = [\,1/2 \;\; 1/2\,]$ is stationary but not limiting distr. because $\bar{\pi} = \bar{\pi} P$ yet $P^{odd} = P$ & $P^{even} = I$.

Matrices with all positive entries are particularly nice – can get to any state from any state with some positive probability.

Def'n: A transition matrix P of a Markov chain is called <u>regular</u> if some power of P has all positive entries ("$P^n > 0$" for some $n \geq 1$)

Example: $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ not regular because

$P^{odd} = P$ & $P^{even} = I$ have zeros

Example: $P = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$

$$P^2 = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$

yes, regular

Theorem 3.2 Limit Theorem for Regular Markov Chains

Let P be the transition matrix for a Markov chain.

Suppose P is regular (P^n has all positive entries for some $n \geq 1$).

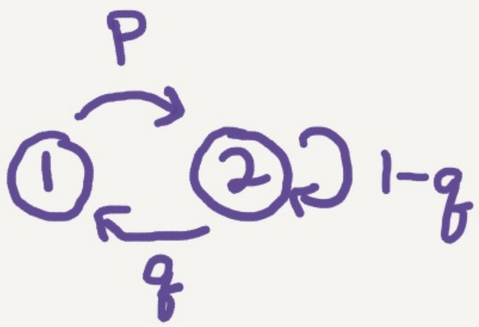
Then P has a limiting distribution which is the unique, positive, stationary distribution $\bar{\pi}$ of the Markov chain: $\bar{\pi} = \bar{\pi} P$

(invariant)

all states get visited

$$\lim_{n \to \infty} P^n = \begin{bmatrix} \bar{\pi} \\ \vdots \\ \bar{\pi} \end{bmatrix}$$

Example: 2-state chain



Transition matrix $P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$ is regular if $p, q > 0$

Solve $\bar{\pi} = \bar{\pi} P$: $[\pi_1 \ \pi_2] = [\pi_1 \ \pi_2] \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$

$$= [\pi_1(1-p) + \pi_2 q \quad \pi_1 p + \pi_2(1-q)]$$

Solve $\pi_1 = \pi_1(1-p) + \pi_2 q$
$\pi_2 = \pi_1 p + \pi_2(1-q)$ $\rightarrow$ $p\pi_1 - q\pi_2 = 0$
$p\pi_1 - q\pi_2 = 0$ $\rightarrow$ $\pi_2 = \frac{p}{q}\pi_1$

Need $\pi_1 + \pi_2 = 1$

$$\Rightarrow \bar{\pi} = \left[\frac{q}{p+q} \quad \frac{p}{p+q} \right]$$

Section 3.3 Classification of states

"Can you find the way to state a?"

In general, we can't expect to have a unique (stationary) invariant prob vector $\bar{\pi}$.

Example: a simple chain with reflecting boundaries has different limits for P^n with n odd vs n even.

Example: a chain with absorbing states has different limits corresponding to each absorbing state (start there, stay there)

Example: has two components, can never get from one to the other, so different limiting distr depending where chain starts

We want to classify these different situations.

Def'n: We say states i & j of a Markov chain <u>communicate</u> with each other ("$i \leftrightarrow j$") if there exists $m, n \geq 0$ such that $P^m_{ij} > 0$ and $P^n_{ji} > 0$. That is, the chain has a positive prob of going from state i to state j in a finite # of steps, and vice versa.

This defines an equivalence relation on the state space

1) Reflexive: $i \leftrightarrow i$ (in 0 steps)
2) Symmetric: $i \leftrightarrow j$ if and only if $j \leftrightarrow i$ (by the def'n)
3) Transitive: if $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$

Equivalence relations partition the space into disjoint sets, which here we will call <u>communication classes</u>.

Def'n: A Markov chain is called <u>irreducible</u> if it has only one communication class, that is, there is a positive prob of getting from every state to every other state in a finite # of steps.

Otherwise, chain is called <u>reducible</u>.

If a chain leaves a communication class eventually with probability 1, then states in that class are called <u>transient</u>.
Otherwise, called <u>recurrent</u> (never leaves that class).

Example: Chain w/ reflecting boundaries

irreducible
$\{1, 2, 3, 4\}$
recurrent

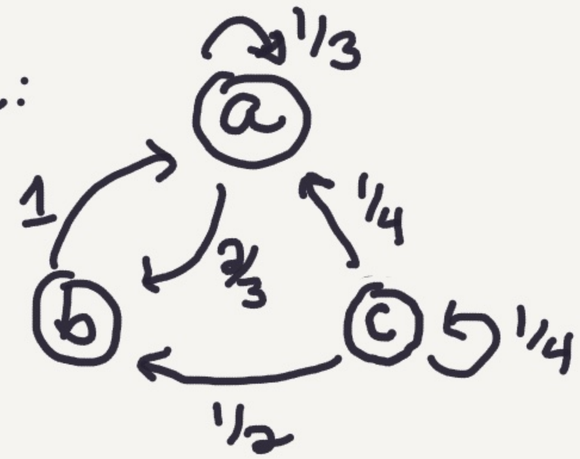
Chain w/ absorbing boundaries

Communication classes: $\{1\}$ $\{5\}$ $\{2, 3, 4\}$

recurrent states

transient

Example:



Communication classes: $\{c\}$ transient

$\{a, b\}$ recurrent

Questions: On average, how many steps until leave state c forever?

How many steps on average to return to a from a, or get to b from a?

Defn: The <u>first passage time</u> to state j is $T_j = \min\{n>0 : X_n = j\}$

(if $X_n \neq j$ for all $n>0$, say $T_j = \infty$)

The prob of ever returning to j is $f_j = \mathbb{P}\{T_j < \infty : X_0 = j\}$

For recurrent states, $f_j = 1$; for transient states, $f_j < 1$.

In example above, $f_a = 1$ $\quad f_b = 1$ $\quad f_c = 1/4$

Whether a state is eventually revisited is related to how often we expect that state to be visited.

Start chain at state $X_0 = i$ and let j be another state.

Let $I_n = \begin{cases} 1 & X_n = j \\ 0 & X_n \neq j \end{cases}$, so $\sum_{n=0}^{\infty} I_n$ = # visits to j

"indicator fn"

$$\mathbb{E}\left[\sum_{n=0}^{\infty} I_n\right] = \sum_{n=0}^{\infty} \mathbb{E}[I_n] = \sum_{n=0}^{\infty} \mathbb{P}\{X_n = j \mid X_0 = i\} = \sum_{n=0}^{\infty} P^n_{ij}$$ (may or may not converge)

(expected # of visits to state j)

Suppose state j is <u>recurrent</u>, so chain started at j eventually revisits j.
Once chain hits j, process begins anew behaving like new chain started at j.
Hence we expect j to be visited infinitely often: $\sum_{n=0}^{\infty} P^n_{jj} = \infty$ diverges.

Now suppose state j is <u>transient</u>.

Prob of hitting j equals f_j and of never hitting f_j is $1-f_j$.

Think of situation as a series of coin flips : $\mathbb{P}\{\text{head}\} = f_j$

$\mathbb{P}\{\text{tail}\} = 1-f_j$

Flip until tail to see how often visit j before leaving forever.

$$\mathbb{E}[\text{flips until tail}] = \underset{\uparrow\ 1^{st}\text{ flip H}}{f_j}\,(\mathbb{E}[\text{flips until tail}]+1) + \underset{\uparrow\ 1^{st}\text{ flip T}}{(1-f_j)}\cdot 1$$

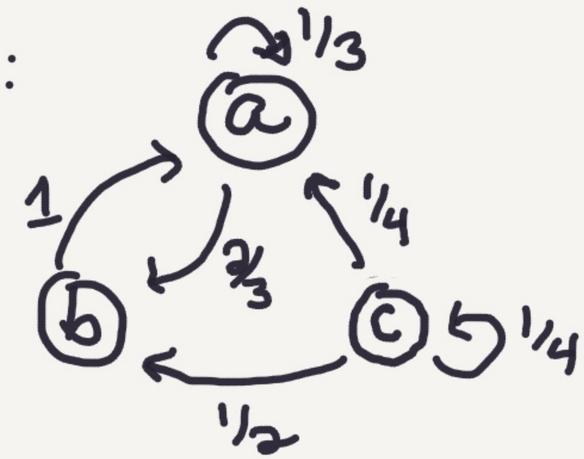
(condition on first flip)

$$= f_j\,\mathbb{E}[\text{flips until tail}] + 1$$

$$\Rightarrow \mathbb{E}[\text{flips until tail}] = \frac{1}{1-f_j}$$

Hence expected # visits to state j is $\sum_{n=0}^{\infty} P^n_{jj} = \frac{1}{1-f_j}$

Return to example:



$$P = \begin{array}{c|ccc} & \text{to } a & b & c \\ \hline \text{from } a & 1/3 & 2/3 & 0 \\ b & 1 & 0 & 0 \\ c & 1/4 & 1/2 & 1/4 \end{array}$$

State a: $P^0_{aa} = 1$

$P^1_{aa} = 1/3$

$P^2_{aa} = 2/3$

$$P^n_{aa} = \frac{1}{3} P^{n-1}_{aa} + \frac{2}{3} P^{n-2}_{aa}$$

a to a in n-1 steps
a to a in n^{th} step

a to a in n-2 steps
a to b then b to a

$\geq 1/3$ by induction

$$\Rightarrow \sum_{n=0}^{\infty} P^n_{aa} = \infty$$

State c: $\sum_{n=0}^{\infty} P^n_{cc} = \sum_{n=0}^{\infty} (1/4)^n = \frac{1}{1-1/4} = \frac{4}{3} = \frac{1}{1-f_c}$, so $f_c = 1/4$

(have to return to c every step)

geometric series formula $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$ for $|r|<1$

Theorem on Recurrence and Transience

For a Markov chain with a finite state space and transition matrix P,

$$\sum_{n=0}^{\infty} P_{jj}^{n} = \infty \iff \text{state } j \text{ is recurrent}$$

$$\sum_{n=0}^{\infty} P_{jj}^{n} < \infty \iff \text{state } j \text{ is transient}$$

<u>Corollary</u>: For finite-state, irreducible Markov chains, all states are recurrent.

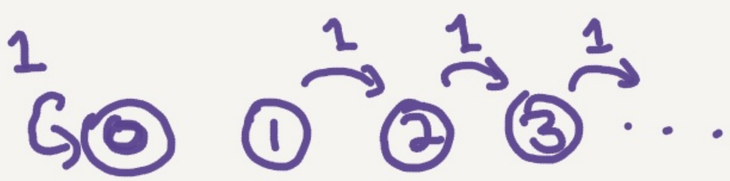
(See proof of Thm 3.3, which shows $\sum_n P_{ij}^n$ converges $\iff$ $\sum_n P_{jj}^n$ converges whenever states i & j are in same communication class)

Defn: A set C of states is called <u>closed</u> if no state outside of C is accessible from any state inside C.

If C is closed, then $P_{ij} = 0$ for all $i \in C$, $j \notin C$.

<u>Lemma 3.5</u>: A communication class with only recurrent states is closed.

A <u>finite</u> communication class is closed <u>only if</u> it consists of recurrent states.

Example: 

(why infinite chains are tricky)

$\{0\}$ recurrent & closed

$\{1, 2, 3, \cdots\}$ transient & closed

<u>Canonical decomposition</u> : partition a finite Markov chain into communication classes R_i of recurrent states and a set T of transient states.

Reorder states to organize transition matrix as

$$P = \begin{array}{c} \\ T \\ R_1 \\ \vdots \\ R_m \end{array} \begin{array}{c} \begin{array}{cccc} T & R_1 & \cdots & R_m \end{array} \\ \begin{bmatrix} * & * & \cdots & * \\ 0 & [\ ] & & 0 \\ \vdots & 0 & & \vdots \\ 0 & 0 & & [\ ] \end{bmatrix} \end{array}$$

square blocks for each R_i
taking powers of P gives
same structure with blocks $[\ ]^n$

This breaks down the Markov chain into components that are easier to analyze for long term behavior – the chain eventually leaves the transient states and remains in one of the recurrent classes.

Section 3.4 Irreducible Markov chains

Recall that a Markov chain is called <u>irreducible</u> if it has only one communication class, that is, there is a positive prob of getting from every state to every other state in a finite # of steps.

<u>Theorem 3.6</u> Limit Theorem for Finite Irreducible Markov chains

Suppose $X_0, X_1, X_2, \cdots$ is a <u>finite</u> <u>irreducible</u> Markov chain.

For each state j, let $\mu_j = \mathbb{E}[T_j \mid X_0 = j]$ expected return time to state j

Then μ_j is finite and there exists a unique stationary distribution $\bar{\pi}$ such that $\pi_j = \frac{1}{\mu_j}$ for all states j.

Furthermore, $\pi_j = \lim_{n\to\infty}\left(\frac{1}{n}\sum_{m=0}^{n-1} P_{ij}^{m}\right)$ for all states i and j

any initial state

prob of going from i to j in m steps

(See rigorous proof in section 3.10 – we'll just get some intuition on why this is true.)

<u>Intuition</u>: π_j = proportion time spent in state j

Given a finite, irreducible MC, start at state j $(X_0 = j)$ and let $T_j = \min\{n \geq 1 : X_n = j\}$ first return time

$\mathbb{E}[T_j \mid X_0 = j] < \infty$ because finite # of states and positive prob of getting between any pair of steps in a finite time

Consider k^{th} return to state j: $T_j^{(1)} + T_j^{(2)} + \cdots + T_j^{(k)}$

(all independent identically distributed RVs)

Law of Large Numbers says $\frac{1}{k}\left(T_j^{(1)} + T_j^{(2)} + \cdots + T_j^{(k)}\right) \approx \mathbb{E}[T_j]$

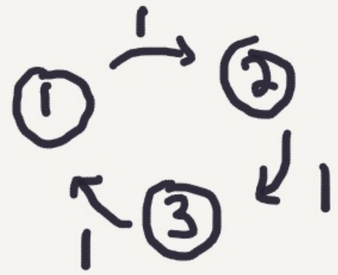
(sample mean $\to$ true mean as sample size $\to \infty$)

So we expect around k visits to state j in $n = k\,\mathbb{E}[T_j]$ steps

We also expect around $n\pi_j$ visits to j after n steps,

so $\underset{\text{visits}}{k} \approx \underbrace{k\,\mathbb{E}[T_j]}_{n \text{ steps}}\pi_j \implies \boxed{\mathbb{E}[T_j] = \frac{1}{\pi_j}}$

Example:



$\pi_1 = 1/3$

$\mathbb{E}[T_1] = 3$

$\boxed{\text{R demo}}$

$$P = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & .4 & .2 & .4 \\ 2 & .6 & 0 & .4 \\ 3 & .2 & .5 & .3 \end{array} \qquad \begin{array}{ccc} \pi_1 & \pi_2 & \pi_3 \\ \end{array}$$

$$\bar{\pi} = [\,.3\overline{78} \quad .2\overline{57} \quad .3\overline{6}\,]$$

Start in state 2 $\Rightarrow$ $\mathbb{E}[T_2] = \frac{1}{\pi_2} \approx 3.88$

Simulation shows spread of return times

For <u>infinite</u> state spaces, can have more complicated situations.

Def'n: State j is called <u>positive recurrent</u> if Markov chain will eventually revisit state j with prob 1 and expected first return time is <u>finite</u>.

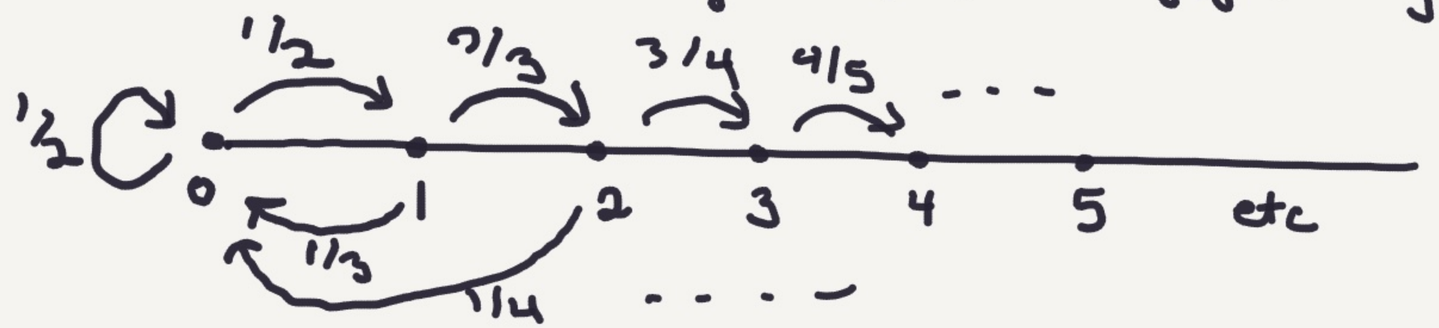
State j is called <u>null recurrent</u> if Markov chain will eventually revisit state j with prob 1 but expected first return time is <u>infinite</u>. (can only occur for infinite state space)

State j is called <u>transient</u> if the prob of returning to j is less than 1.

All states in the same communication class will share the same classification.

Example: State space $\{0, 1, 2, \cdots\}$

Transition prob $P_{j0} = \frac{1}{j+2}$ $\quad P_{j,j+1} = \frac{j+1}{j+2}$ for $j \geq 0$



First return time to state 0: look at prob of <u>never</u> returning

$$\mathbb{P}\{T_0 = \infty \mid X_0 = 0\} = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdots$$

$$= \prod_{j=0}^{\infty} \frac{j+1}{j+2} = \lim_{j \to \infty} \frac{1}{j+2} = 0$$

So the prob of returning to state 0 is 1.

The prob of returning to state 0 in exactly k steps is

$$\mathbb{P}\{T_0 = k \mid X_0 = 0\} = \underbrace{\frac{1}{2} \cdot \frac{2}{3} \cdots \frac{k-1}{k}}_{\text{march out to state } j+1 = k-1} \cdot \underbrace{\frac{1}{(k-1)+2}}_{\text{return to } 0} = \frac{1}{k(k+1)}$$

Expected first return time to state 0: $\mathbb{E}[T_0 \mid X_0 = 0] = \sum_{k=1}^{\infty} k \mathbb{P}\{T_0 = k \mid X_0 = 0\}$

$$= \sum_{k=1}^{\infty} \frac{k}{k(k+1)} = \sum_{k=1}^{\infty} \frac{1}{k+1}$$

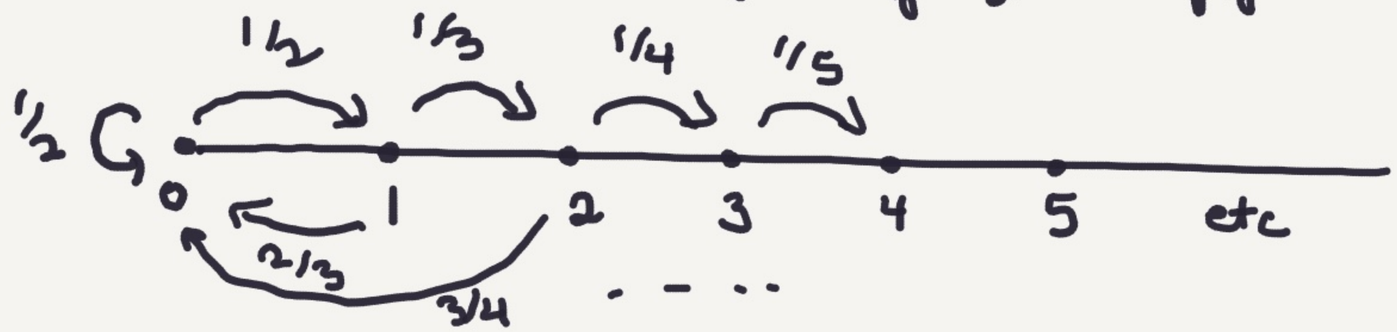
Harmonic series, diverges

Hence state 0 is null recurrent.

Irreducible MC, so all states are <u>null recurrent.</u>

Example: State space $\{0,1,2,\cdots\}$

Transition prob $P_{j0} = \frac{j+1}{j+2}$ $\quad P_{j,j+1} = \frac{1}{j+2}$



$$\mathbb{P}\{T_0 = \infty \mid X_0 = 0\} = \frac{1}{2}\cdot\frac{1}{3}\cdot\frac{1}{4}\cdots = \prod_{j=0}^{\infty}\frac{1}{j+2} = 0$$

so returns to 0 with prob 1

$$\mathbb{P}\{T_0 = k \mid X_0 = 0\} = \frac{1}{2}\cdot\frac{1}{3}\cdots\underbrace{\frac{1}{(k-2)+2}}_{\text{march out to state } j+1=k-1}\cdot\underbrace{\frac{k}{k+1}}_{\text{return to 0 from } j=k-1} = \frac{k}{(k+1)!}$$

Expected first return time $\mathbb{E}[T_0 \mid X_0 = 0] = \sum_{k=1}^{\infty} k\cdot\frac{k}{(k+1)!} < \infty$

by ratio test

Hence state 0 is positive recurrent,

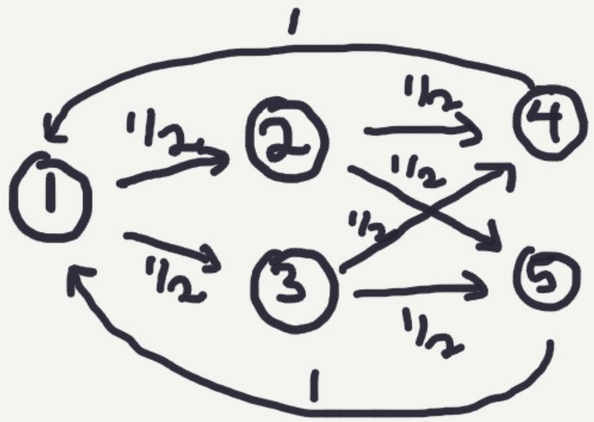
and so all states are <u>positive recurrent</u>

(irreducible MC, single communication class)

Section 3.5 Periodicity (in a stochastic sense)

Def'n: The <u>period</u> $d = d(i)$ of a state i of a MC is the greatest common divisor of $J_i = \{n \geq 0 : P^n_{ii} > 0\}$

Example:



$\boxed{\text{R demo}}$

P^{3m} same for $m \geq 1$, all positive diagonal entries

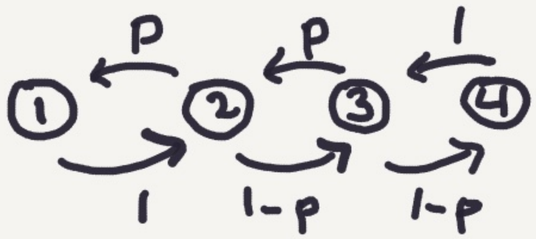
$J_i = \{0, 3, 6, 9, \cdots\}$ for all states i

Period $d = d(i) = 3$ for all i

Lemma 3.7 All states of an irreducible MC have the same period. (See proof in book.)

Def'n: We call an irreducible MC <u>aperiodic</u> if period equals 1.

Example: 4-state reflecting boundaries



$\boxed{\text{R demo}}$

P^{2m} has positive diagonal entries for $m \geq 1$ but odd powers of P have zeros

$J_i = \{0, 2, 4, 6, \cdots\}$ and $d = 2$ for all states

Theorem: If P is the transition matrix for an irreducible, finite, aperiodic Markov chain, then there exists a unique stationary vector $\bar{\pi}$ satisfying $\bar{\pi} P = \bar{\pi}$. If $\bar{\alpha}$ is any initial prob vector, then $\lim_{n\to\infty} \bar{\alpha} P^n = \bar{\pi}$. Moreover, $\pi_i > 0$ for all states i.

visit every state over long run

initial state doesn't matter in long run

Proof: We can apply our prior thm on irreducible MCs if we can show $P^n > 0$ for some $n \geq 1$ (P is regular).

Take any states i and j. Since P is irreducible, $P_{ij}^{m_{ij}} > 0$ for some $m_{ij} \geq 1$

Because P is aperiodic, there exists M_i such that for all $n \geq M_i$, $P_{ii}^n > 0$

(number theory result)

Hence for all $n \geq M_i$, $P_{ij}^{n+m_{ij}} \geq P_{ii}^n P_{ij}^{m_{ij}} > 0$

(one way to get from i to j in $n + m_{ij}$ steps)

Let $M = \max\{M_i + m_{ij} : \text{all states } i, j\}$

Then $P_{ij}^n > 0$ for all $n \geq M$ and states i, j.

That is, $P^M > 0$, so the prior theorem applies. $\square$

What about reducible & periodic chains?. What can we say?

Suppose P is reducible with recurrent classes $R_1, \cdots, R_r$
(once in one of these classes, stays there)

and transient classes $T_1, \cdots, T_s$
(eventually leaves)

If the recurrent classes are aperiodic (treated as their own MC), then we have a stationary vector π^k for each R^k
(can set $\pi^k_i = 0$ for $i \notin R^k$)

i^{th} component of vector π^k

Then $\lim_{n\to\infty} P^n_{ij} = \begin{cases} \pi^k_j & \text{for } i,j \in R^k \\ 0 & \text{o.w.} \end{cases}$

$\lim_{n\to\infty} P^n_{ij} = 0$ for all transient states

Note that if we have more than one recurrent class, where the chain starts will be important for which recurrent class it eventually ends up in.

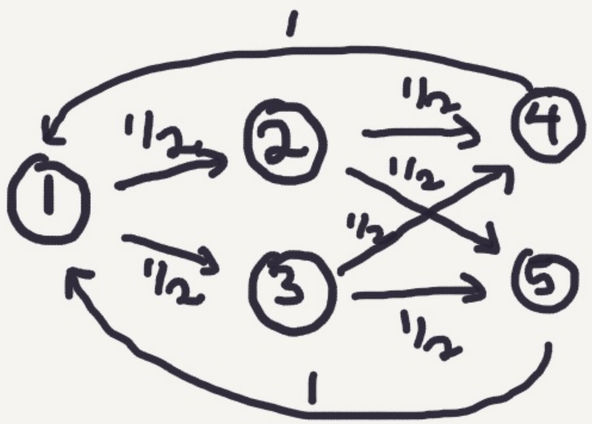
Now suppose the Markov chain is irreducible but periodic, $d > 1$

In this case, the chain splits into d sets of states $A_1, \cdots, A_d$

such that

$$\begin{tikzcd} A_1 \arrow[r] & A_2 \arrow[dr] & \\ A_d \arrow[u] & \cdots \arrow[l] & A_3 \arrow[l] \end{tikzcd}$$

Consider previous example:



$\boxed{\text{R demo}}$

Eigenvalues of P are

$$\lambda = \underbrace{1, \ -\tfrac{1}{2} \pm \tfrac{\sqrt{3}}{2} i}_{3^{rd} \text{ roots of unity}}, 0, 0$$

Eigenvector for $\lambda = 1$ is $[1/3 \ \ 1/6 \ \ 1/6 \ \ 1/6 \ \ 1/6]$

A bit of linear algebra will reveal what's happening for period d chains.

Most transition matrices are "diagonalizable" (except for rare cases)

The idea is that an $n \times n$ matrix P has n eigenvalues (can be duplication) associated with n linearly independent eigenvectors.

Let $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ diagonal matrix of eigenvalues

Q = matrix with associated right eigenvectors $P\bar{w}_k = \lambda_k \bar{w}_k$

$= [\bar{w}_1 \ \bar{w}_2 \ \cdots \ \bar{w}_n]$

Then $Q^{-1}PQ = Q^{-1}[P\bar{w}_1 \cdots P\bar{w}_n] = Q^{-1}[\lambda_1\bar{w}_1 \cdots \lambda_n\bar{w}_n]$

$$= Q^{-1}\underbrace{[\bar{w}_1 \cdots \bar{w}_n]}_{Q}\underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}}_{D}$$

$$= D$$

Hence $D = Q^{-1}PQ$,

so $DQ^{-1} = Q^{-1}P$

Let $\bar{v}_1, \cdots, \bar{v}_n$ be the rows of $Q^{-1} = \begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix}$

$$DQ^{-1} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}\begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\bar{v}_1 \\ \vdots \\ \lambda_n\bar{v}_n \end{bmatrix} \Rightarrow \bar{v}_k P = \lambda_k P, \ k = 1, \cdots, n$$

$$\| $$

$$Q^{-1}P = \begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix} P = \begin{bmatrix} \bar{v}_1 P \\ \vdots \\ \bar{v}_n P \end{bmatrix}$$

Rows of Q^{-1} are the left eigenvectors of P.

One of the nice properties of diagonalization is taking powers of the matrix:

$$D = Q^{-1}PQ \Rightarrow P = QDQ^{-1} \Rightarrow P^2 = (QDQ^{-1})(QDQ^{-1})$$
$$= QD^2Q^{-1}$$
$$= \begin{bmatrix} \bar{w}_1 & \cdots & \bar{w}_n \end{bmatrix} \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{bmatrix} \begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix}$$

Generally, $P^m = QD^mQ^{-1}$

$$= \begin{bmatrix} \bar{w}_1 & \cdots & \bar{w}_n \end{bmatrix} \begin{bmatrix} \lambda_1^m & & \\ & \ddots & \\ & & \lambda_n^m \end{bmatrix} \begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix}$$
$$= \lambda_1^m \bar{w}_1 \bar{v}_1 + \cdots + \lambda_n^m \bar{w}_n \bar{v}_n$$

If we knew more about $\lambda_1, \cdots, \lambda_n$, we could see what happens as the power m grows large (long term behavior).

Perron-Frobenius Theorem for Irreducible Periodic MCs

Let P be the transition matrix for a finite, irreducible MC with period d. Then the following hold:

1) P has a single eigenvalue $\lambda = 1$ with right eigenvector $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ and left eigenvector with all positive components.

2) P has d eigenvalues that are d^{th} roots of unity:

   (Let $\omega = e^{2\pi i/d} = \cos(2\pi/d) + i\sin(2\pi/d)$, so $\omega^d = 1$ & $|\omega| = 1$
   then d^{th} roots of unity are $1, \omega, \omega^2, \dots, \omega^{d-1}$)

3) Remaining eigenvalues have absolute value less than 1.

   $\hookrightarrow \lambda_{d+1}^m, \dots, \lambda_n^m \to 0$ as $m \to \infty$

This means that $P^m \approx \begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_1 \end{bmatrix} + \omega^m \bar{w}_2 \bar{v}_2 + \omega^{2m} \bar{w}_3 \bar{v}_3 + \dots + \omega^{(d-1)m} \bar{w}_{d-1} \bar{v}_{d-1}$

(why we get d alternating matrices for large powers)

We can also think about the average time spent in each state by averaging across the cycle: Proportion time spent in state j if start at state i $= \lim_{m\to\infty} (P_{ij}^{m+1} + \dots + P_{ij}^{m+d})/d$

R demo